



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 1, 2023 – 02:00 PM JST

PDB ID : 5IG7  
Title : Crystal structure of anti-gliadin 1002-1E01 Fab fragment in complex of peptide PLQPQQPFP  
Authors : Snir, O.; Chen, X.; Gidoni, M.; du Pre, M.F.; Zhao, Y.; Steinsbo, O.; Lundin, K.E.; Yaari, G.; Sollid, L.M.  
Deposited on : 2016-02-27  
Resolution : 1.97 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

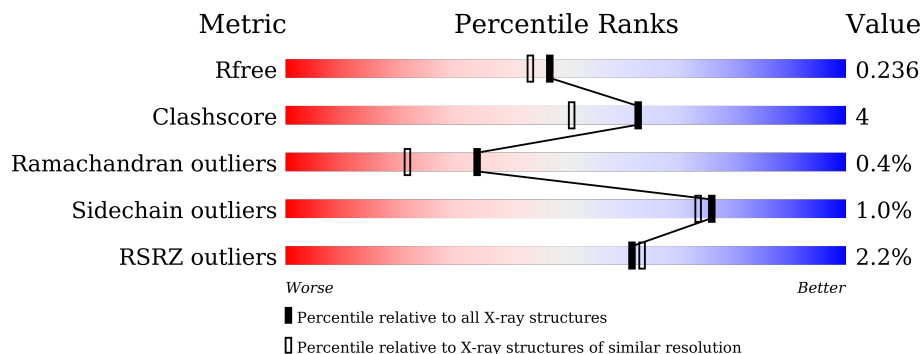
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.97 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	227	 91% 5%
1	D	227	 88% 7% ..
1	G	227	 87% 8% 5%
1	J	227	 89% 7% .
2	B	216	 90% 8% ..
2	E	216	 89% 9% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	216	 5% 88% 11%
2	K	216	 2% 88% 9%
3	C	9	 78% 22%
3	F	9	 44% 56%
3	I	9	 78% 22%
3	L	9	 56% 44%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 13917 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called 1E01 Fab fragment heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	216	Total 1593	C 1000	N 267	O 318	S 8	0	0	0
1	D	217	Total 1602	C 1006	N 269	O 319	S 8	0	0	0
1	G	216	Total 1593	C 1000	N 267	O 318	S 8	0	0	0
1	J	219	Total 1614	C 1012	N 271	O 323	S 8	0	0	0

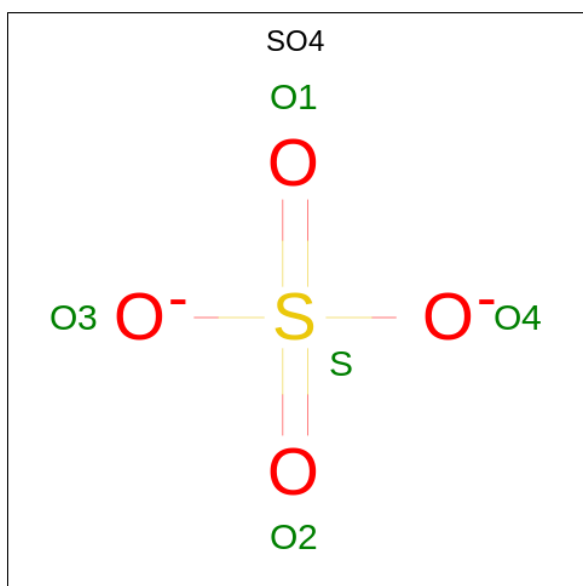
- Molecule 2 is a protein called 1E01 Fab fragment light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	213	Total 1593	C 993	N 270	O 325	S 5	0	0	0
2	E	213	Total 1593	C 993	N 270	O 325	S 5	0	0	0
2	H	213	Total 1593	C 993	N 270	O 325	S 5	0	0	0
2	K	213	Total 1593	C 993	N 270	O 325	S 5	0	0	0

- Molecule 3 is a protein called peptide PRO-LEU-GLN-PRO-GLN-GLN-PRO-PHE-PRO.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	9	Total 74	C 50	N 12	O 12	0	0	0
3	F	9	Total 74	C 50	N 12	O 12	0	0	0
3	I	9	Total 74	C 50	N 12	O 12	0	0	0
3	L	9	Total 74	C 50	N 12	O 12	0	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	E	1	Total	O	S	0	0
			5	4	1		
4	E	1	Total	O	S	0	0
			5	4	1		
4	G	1	Total	O	S	0	0
			5	4	1		
4	J	1	Total	O	S	0	0
			5	4	1		
4	J	1	Total	O	S	0	0
			5	4	1		
4	K	1	Total	O	S	0	0
			5	4	1		
4	K	1	Total	O	S	0	0
			5	4	1		
4	K	1	Total	O	S	0	0
			5	4	1		

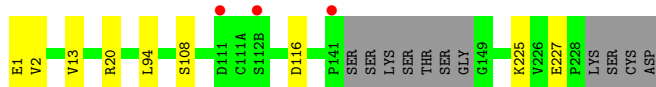
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	140	Total O 140 140	0	0
5	B	81	Total O 81 81	0	0
5	C	4	Total O 4 4	0	0
5	D	112	Total O 112 112	0	0
5	E	72	Total O 72 72	0	0
5	F	5	Total O 5 5	0	0
5	G	115	Total O 115 115	0	0
5	H	54	Total O 54 54	0	0
5	I	2	Total O 2 2	0	0
5	J	126	Total O 126 126	0	0
5	K	79	Total O 79 79	0	0
5	L	2	Total O 2 2	0	0

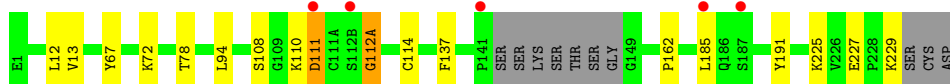
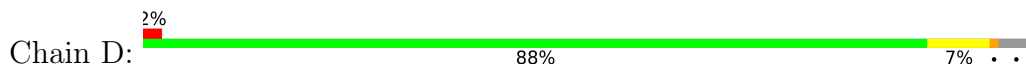
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

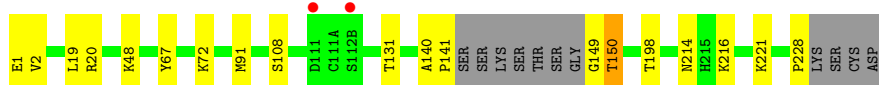
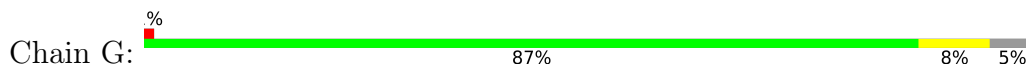
- Molecule 1: 1E01 Fab fragment heavy chain



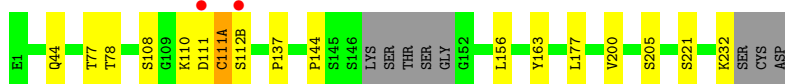
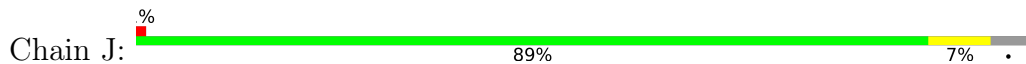
- Molecule 1: 1E01 Fab fragment heavy chain



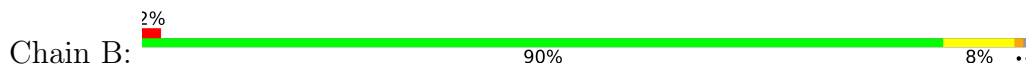
- Molecule 1: 1E01 Fab fragment heavy chain



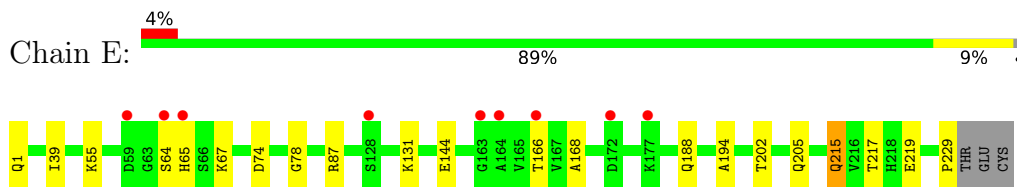
- Molecule 1: 1E01 Fab fragment heavy chain



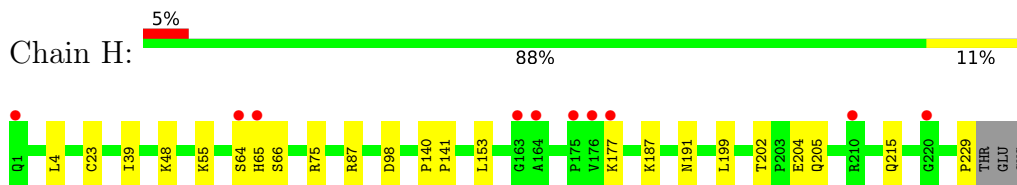
- Molecule 2: 1E01 Fab fragment light chain



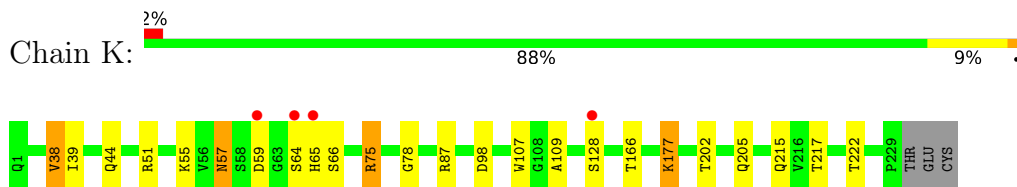
- Molecule 2: 1E01 Fab fragment light chain



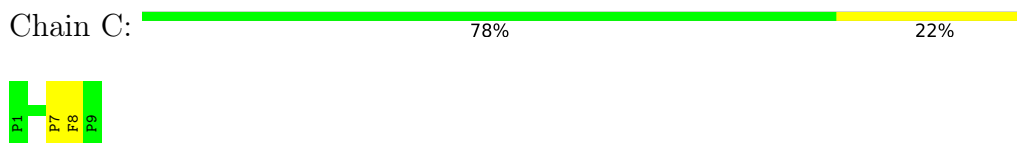
- Molecule 2: 1E01 Fab fragment light chain



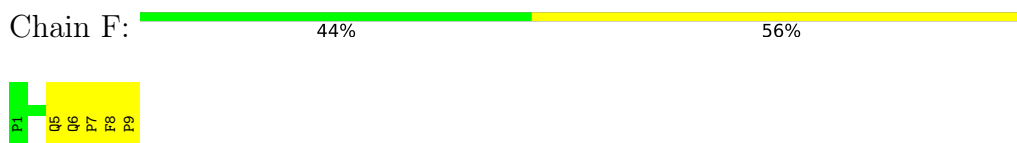
- Molecule 2: 1E01 Fab fragment light chain



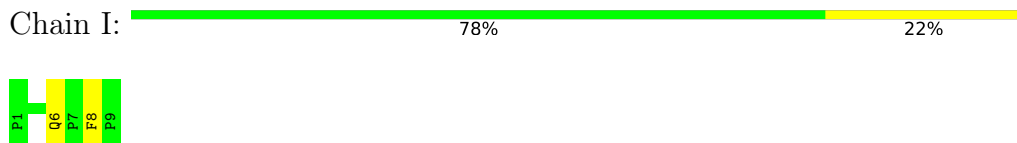
- Molecule 3: peptide PRO-LEU-GLN-PRO-GLN-GLN-PRO-PHE-PRO



- Molecule 3: peptide PRO-LEU-GLN-PRO-GLN-GLN-PRO-PHE-PRO



- Molecule 3: peptide PRO-LEU-GLN-PRO-GLN-GLN-PRO-PHE-PRO



- Molecule 3: peptide PRO-LEU-GLN-PRO-GLN-GLN-PRO-PHE-PRO





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.90Å 112.67Å 104.95Å 90.00° 103.47° 90.00°	Depositor
Resolution (Å)	55.87 – 1.97 55.87 – 1.97	Depositor EDS
% Data completeness (in resolution range)	99.8 (55.87-1.97) 92.7 (55.87-1.97)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.15 (at 1.97Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155)	Depositor
R, $R_{free}$	0.190 , 0.236 0.189 , 0.236	Depositor DCC
$R_{free}$ test set	6375 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.1	Xtrriage
Anisotropy	0.426	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 39.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13917	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.65% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.41	0/1629	0.59	0/2215
1	D	0.41	0/1638	0.61	1/2226 (0.0%)
1	G	0.40	0/1629	0.58	0/2215
1	J	0.42	0/1650	0.59	0/2242
2	B	0.38	0/1630	0.58	0/2219
2	E	0.36	0/1630	0.57	0/2219
2	H	0.34	0/1630	0.55	0/2219
2	K	0.36	0/1630	0.58	0/2219
3	C	0.36	0/78	0.40	0/107
3	F	0.38	0/78	0.42	0/107
3	I	0.39	0/78	0.64	0/107
3	L	0.41	0/78	0.51	0/107
All	All	0.39	0/13378	0.58	1/18202 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	112(A)	GLY	N-CA-C	-5.89	98.38	113.10

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1593	0	1550	8	0
1	D	1602	0	1563	14	0
1	G	1593	0	1550	13	1
1	J	1614	0	1573	13	0
2	B	1593	0	1546	13	0
2	E	1593	0	1546	16	0
2	H	1593	0	1546	15	0
2	K	1593	0	1546	21	1
3	C	74	0	74	2	0
3	F	74	0	74	9	0
3	I	74	0	74	2	0
3	L	74	0	74	4	0
4	B	5	0	0	0	0
4	D	5	0	0	0	0
4	E	10	0	0	0	0
4	G	5	0	0	0	0
4	J	10	0	0	0	0
4	K	20	0	0	0	0
5	A	140	0	0	1	0
5	B	81	0	0	2	0
5	C	4	0	0	0	0
5	D	112	0	0	3	1
5	E	72	0	0	2	1
5	F	5	0	0	0	0
5	G	115	0	0	2	0
5	H	54	0	0	1	0
5	I	2	0	0	1	0
5	J	126	0	0	1	0
5	K	79	0	0	2	0
5	L	2	0	0	0	0
All	All	13917	0	12716	112	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (112) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1:GLU:HG2	1:A:2:VAL:H	1.36	0.90
2:E:188:GLN:HE21	2:E:194:ALA:HB2	1.35	0.89
3:F:6:GLN:HE22	3:F:8:PHE:C	1.75	0.88
2:K:75:ARG:NH2	2:K:98:ASP:OD2	2.09	0.85

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:1:GLU:HG2	1:G:2:VAL:H	1.40	0.84
3:F:6:GLN:HE22	3:F:9:PRO:N	1.76	0.83
2:H:75:ARG:NH2	2:H:98:ASP:OD2	2.15	0.78
3:I:6:GLN:HE21	3:I:8:PHE:H	1.29	0.78
3:F:6:GLN:OE1	3:F:8:PHE:N	2.17	0.77
2:E:64:SER:O	2:E:65:HIS:ND1	2.17	0.77
3:F:6:GLN:CD	3:F:8:PHE:H	1.90	0.74
2:H:64:SER:O	2:H:65:HIS:ND1	2.21	0.72
2:K:64:SER:O	2:K:65:HIS:ND1	2.21	0.72
2:K:57:ASN:HD22	2:K:59:ASP:H	1.39	0.71
2:K:177:LYS:HD3	2:K:177:LYS:H	1.56	0.70
2:E:229:PRO:O	5:E:401:HOH:O	2.11	0.68
2:K:65:HIS:O	5:K:401:HOH:O	2.10	0.68
1:D:110:LYS:O	5:D:401:HOH:O	2.12	0.68
1:J:144:PRO:O	1:J:232:LYS:NZ	2.27	0.67
1:A:1:GLU:CG	1:A:2:VAL:H	2.07	0.67
2:B:229:PRO:O	5:B:401:HOH:O	2.12	0.67
1:J:44:GLN:HE22	2:K:44:GLN:HE22	1.44	0.66
2:E:74:ASP:OD1	5:E:402:HOH:O	2.15	0.63
3:I:6:GLN:OE1	5:I:101:HOH:O	2.15	0.63
1:A:20:ARG:NE	5:A:301:HOH:O	2.32	0.62
2:H:202:THR:H	2:H:205:GLN:HE21	1.46	0.61
1:D:13:VAL:HG11	1:D:94:LEU:HD13	1.82	0.61
2:H:202:THR:OG1	2:H:205:GLN:HG3	2.00	0.61
5:D:423:HOH:O	1:J:77:THR:HG21	2.00	0.60
2:K:202:THR:OG1	2:K:205:GLN:HG3	2.00	0.60
2:B:51:ARG:NH1	5:B:403:HOH:O	2.34	0.60
2:K:109:ALA:HA	3:L:4:PRO:HG3	1.82	0.60
2:B:202:THR:OG1	2:B:205:GLN:HG3	2.02	0.60
1:J:78:THR:HG21	5:J:486:HOH:O	2.03	0.59
3:F:6:GLN:NE2	3:F:8:PHE:C	2.54	0.58
2:H:153:LEU:HD12	2:H:199:LEU:HD23	1.86	0.57
1:D:78:THR:HG21	5:D:472:HOH:O	2.03	0.57
2:E:65:HIS:CE1	2:E:78:GLY:H	2.22	0.57
1:J:144:PRO:HG3	1:J:156:LEU:HB3	1.87	0.56
2:E:202:THR:OG1	2:E:205:GLN:HG3	2.05	0.56
1:D:112(A):GLY:HA2	1:D:114:CYS:H	1.70	0.56
2:H:39:ILE:HG12	2:H:87:ARG:HG3	1.88	0.56
1:A:108:SER:OG	2:B:55:LYS:NZ	2.39	0.55
1:G:214:ASN:CG	1:G:221:LYS:HE2	2.26	0.55
1:A:13:VAL:HG11	1:A:94:LEU:HD13	1.88	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:J:111:ASP:HB3	3:L:5:GLN:HE21	1.71	0.55
2:K:217:THR:HG22	2:K:222:THR:OG1	2.06	0.55
2:B:39:ILE:HG12	2:B:87:ARG:HG3	1.89	0.55
2:E:1:GLN:OE1	2:E:1:GLN:N	2.37	0.55
1:D:225:LYS:HE2	1:D:227:GLU:HG2	1.88	0.55
1:D:185:LEU:HD23	1:D:191:TYR:CZ	2.42	0.54
3:F:6:GLN:HG2	3:F:7:PRO:HD2	1.89	0.54
2:E:67:LYS:NZ	2:E:74:ASP:HA	2.23	0.54
2:K:57:ASN:ND2	2:K:59:ASP:H	2.06	0.53
1:A:225:LYS:HE3	1:A:227:GLU:HG2	1.90	0.53
1:G:1:GLU:CG	1:G:2:VAL:H	2.15	0.52
1:D:111:ASP:HB2	3:F:5:GLN:HB3	1.92	0.52
2:K:75:ARG:HH22	2:K:98:ASP:CG	2.12	0.52
2:K:65:HIS:CE1	2:K:78:GLY:H	2.28	0.51
2:K:51:ARG:NH1	5:K:404:HOH:O	2.38	0.51
2:K:166:THR:HB	2:K:217:THR:OG1	2.11	0.51
1:J:110:LYS:O	1:J:111:ASP:CG	2.49	0.51
1:D:67:TYR:HB2	1:D:72:LYS:HD2	1.93	0.50
2:H:229:PRO:O	5:H:301:HOH:O	2.17	0.50
1:D:137:PHE:CE1	2:E:144:GLU:HG2	2.46	0.49
1:G:67:TYR:HB2	1:G:72:LYS:HD2	1.93	0.49
2:E:168:ALA:O	2:E:215:GLN:HB3	2.13	0.49
1:D:137:PHE:CD1	2:E:144:GLU:HG2	2.48	0.48
1:D:111:ASP:HB2	3:F:5:GLN:HE21	1.79	0.47
1:A:1:GLU:HG2	1:A:2:VAL:N	2.17	0.47
2:K:57:ASN:HD21	2:K:59:ASP:HB2	1.80	0.47
1:D:185:LEU:HD23	1:D:191:TYR:CE2	2.51	0.46
2:K:38:VAL:HG13	3:L:3:GLN:OE1	2.16	0.46
2:H:75:ARG:HH22	2:H:98:ASP:CG	2.20	0.46
2:B:138:LEU:HD13	2:B:227:VAL:HG13	1.97	0.46
2:E:39:ILE:HG12	2:E:87:ARG:HG3	1.98	0.46
1:J:137:PRO:HB3	1:J:163:TYR:HB3	1.97	0.45
2:H:140:PRO:HA	2:H:153:LEU:HD23	1.97	0.45
2:H:141:PRO:HD3	2:H:153:LEU:CD2	2.46	0.45
2:H:187:LYS:HE2	2:H:191:ASN:OD1	2.17	0.45
1:D:108:SER:OG	2:E:55:LYS:NZ	2.49	0.44
1:A:116:ASP:HB2	2:B:52:TYR:CD1	2.53	0.44
3:F:6:GLN:OE1	3:F:8:PHE:CA	2.66	0.44
2:K:107:TRP:HB2	3:L:7:PRO:HB3	2.00	0.44
2:E:166:THR:OG1	2:E:217:THR:OG1	2.34	0.44
2:K:39:ILE:HG12	2:K:87:ARG:HG3	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:131:THR:HG21	1:J:221:SER:O	2.18	0.43
2:B:48:LYS:HB2	2:B:48:LYS:HE2	1.83	0.43
1:J:177:LEU:HD21	1:J:200:VAL:HG21	2.01	0.43
1:J:111(A):CYS:O	1:J:112(B):SER:HB3	2.19	0.43
2:E:65:HIS:HE1	2:E:78:GLY:O	2.02	0.43
2:B:38:VAL:HG22	2:B:57:ASN:HB3	2.02	0.42
2:E:131:LYS:HD2	2:E:219:GLU:HG3	2.00	0.42
1:G:108:SER:OG	2:H:55:LYS:NZ	2.52	0.42
2:K:65:HIS:HE1	2:K:78:GLY:O	2.03	0.42
2:B:202:THR:OG1	2:B:204:GLU:HG2	2.19	0.42
1:J:108:SER:OG	2:K:55:LYS:NZ	2.52	0.42
1:G:20:ARG:NH1	5:G:401:HOH:O	2.15	0.41
2:B:107:TRP:HB2	3:C:7:PRO:HB3	2.03	0.41
1:G:140:ALA:HB1	1:G:141:PRO:HA	2.03	0.41
1:G:216:LYS:NZ	5:G:406:HOH:O	2.52	0.41
1:J:44:GLN:HE22	2:K:44:GLN:NE2	2.12	0.41
1:G:19:LEU:HB3	1:G:91:MET:HE3	2.03	0.41
1:G:141:PRO:HA	1:G:228:PRO:HA	2.03	0.41
2:B:163:GLY:O	2:B:185:PRO:HG2	2.20	0.41
1:D:12:LEU:HB2	1:D:162:PRO:HG3	2.02	0.41
2:H:4:LEU:HD22	2:H:23:CYS:SG	2.60	0.41
2:H:141:PRO:HD3	2:H:153:LEU:HD23	2.02	0.41
1:G:150:THR:HG23	1:G:198:THR:HG23	2.02	0.41
1:G:214:ASN:CB	1:G:221:LYS:HE2	2.51	0.40
2:H:75:ARG:H	2:H:75:ARG:HG2	1.59	0.40
2:B:116:LEU:HD13	3:C:8:PHE:CZ	2.56	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:D:497:HOH:O	5:E:411:HOH:O[2_1158]	2.11	0.09
1:G:149:GLY:N	2:K:128:SER:OG[1_655]	2.15	0.05

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries

of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	212/227 (93%)	208 (98%)	4 (2%)	0	100	100
1	D	213/227 (94%)	207 (97%)	5 (2%)	1 (0%)	29	16
1	G	212/227 (93%)	208 (98%)	4 (2%)	0	100	100
1	J	215/227 (95%)	211 (98%)	4 (2%)	0	100	100
2	B	211/216 (98%)	201 (95%)	9 (4%)	1 (0%)	29	16
2	E	211/216 (98%)	203 (96%)	7 (3%)	1 (0%)	29	16
2	H	211/216 (98%)	203 (96%)	6 (3%)	2 (1%)	17	8
2	K	211/216 (98%)	202 (96%)	7 (3%)	2 (1%)	17	8
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	7 (100%)	0	0	100	100
3	I	7/9 (78%)	7 (100%)	0	0	100	100
3	L	7/9 (78%)	7 (100%)	0	0	100	100
All	All	1724/1808 (95%)	1670 (97%)	47 (3%)	7 (0%)	34	22

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	215	GLN
2	H	66	SER
2	K	215	GLN
2	B	66	SER
2	K	66	SER
2	H	215	GLN
1	D	111	ASP

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	178/188 (95%)	178 (100%)	0	100	100
1	D	179/188 (95%)	178 (99%)	1 (1%)	86	85
1	G	178/188 (95%)	176 (99%)	2 (1%)	73	70
1	J	181/188 (96%)	179 (99%)	2 (1%)	73	70
2	B	179/182 (98%)	177 (99%)	2 (1%)	73	70
2	E	179/182 (98%)	179 (100%)	0	100	100
2	H	179/182 (98%)	176 (98%)	3 (2%)	60	53
2	K	179/182 (98%)	175 (98%)	4 (2%)	52	46
3	C	9/9 (100%)	9 (100%)	0	100	100
3	F	9/9 (100%)	9 (100%)	0	100	100
3	I	9/9 (100%)	9 (100%)	0	100	100
3	L	9/9 (100%)	9 (100%)	0	100	100
All	All	1468/1516 (97%)	1454 (99%)	14 (1%)	76	73

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	48	LYS
2	B	138	LEU
1	D	229	LYS
1	G	48	LYS
1	G	150	THR
2	H	48	LYS
2	H	177	LYS
2	H	204	GLU
1	J	111(A)	CYS
1	J	205	SER
2	K	38	VAL
2	K	57	ASN
2	K	75	ARG
2	K	177	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	214	ASN
2	B	147	GLN
3	C	5	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	D	179	HIS
2	E	149	ASN
2	E	188	GLN
3	F	5	GLN
2	H	147	GLN
2	H	205	GLN
3	I	6	GLN
1	J	44	GLN
1	J	217	ASN
2	K	57	ASN
3	L	5	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

11 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	SO4	G	301	-	4,4,4	0.14	0	6,6,6	0.16	0
4	SO4	K	301	-	4,4,4	0.14	0	6,6,6	0.10	0
4	SO4	K	304	-	4,4,4	0.13	0	6,6,6	0.06	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	D	301	-	4,4,4	0.15	0	6,6,6	0.11	0
4	SO4	K	302	-	4,4,4	0.15	0	6,6,6	0.04	0
4	SO4	E	301	-	4,4,4	0.14	0	6,6,6	0.12	0
4	SO4	K	303	-	4,4,4	0.14	0	6,6,6	0.09	0
4	SO4	E	302	-	4,4,4	0.15	0	6,6,6	0.08	0
4	SO4	J	302	-	4,4,4	0.12	0	6,6,6	0.16	0
4	SO4	B	301	-	4,4,4	0.14	0	6,6,6	0.08	0
4	SO4	J	301	-	4,4,4	0.16	0	6,6,6	0.20	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	216/227 (95%)	-0.13	3 (1%) 75 77	24, 34, 60, 87	0
1	D	217/227 (95%)	0.00	5 (2%) 60 62	24, 37, 63, 83	0
1	G	216/227 (95%)	-0.10	2 (0%) 84 85	28, 37, 61, 89	0
1	J	219/227 (96%)	-0.10	2 (0%) 84 85	25, 37, 62, 81	0
2	B	213/216 (98%)	-0.07	4 (1%) 66 68	26, 44, 65, 87	0
2	E	213/216 (98%)	0.19	9 (4%) 36 38	26, 46, 74, 89	0
2	H	213/216 (98%)	0.22	10 (4%) 31 33	32, 53, 71, 83	0
2	K	213/216 (98%)	-0.03	4 (1%) 66 68	27, 44, 65, 83	0
3	C	9/9 (100%)	0.03	0 100 100	36, 51, 71, 76	0
3	F	9/9 (100%)	-0.17	0 100 100	33, 39, 50, 58	0
3	I	9/9 (100%)	0.26	0 100 100	37, 48, 88, 93	0
3	L	9/9 (100%)	0.02	0 100 100	28, 39, 55, 69	0
All	All	1756/1808 (97%)	-0.00	39 (2%) 62 63	24, 40, 67, 93	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	111	ASP	5.6
2	B	128	SER	5.4
1	J	112(B)	SER	4.4
2	H	164	ALA	3.8
1	A	111	ASP	3.8
2	B	65	HIS	3.8
2	B	64	SER	3.5
1	G	112(B)	SER	3.5
2	E	177	LYS	3.3
2	E	64	SER	3.3
2	H	65	HIS	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	E	163	GLY	3.0
1	D	187	SER	3.0
2	H	64	SER	2.8
2	E	59	ASP	2.8
2	K	64	SER	2.7
1	J	111	ASP	2.7
2	H	175	PRO	2.6
2	B	59	ASP	2.6
2	K	128	SER	2.5
2	K	65	HIS	2.5
1	D	112(B)	SER	2.4
2	E	65	HIS	2.4
2	H	163	GLY	2.4
1	D	185	LEU	2.3
2	E	128	SER	2.3
2	H	1	GLN	2.2
1	A	141	PRO	2.2
2	H	177	LYS	2.2
2	E	164	ALA	2.2
1	D	141	PRO	2.1
1	D	111	ASP	2.1
2	E	166	THR	2.1
2	E	172	ASP	2.1
2	H	176	VAL	2.1
2	H	210	ARG	2.0
2	H	220	GLY	2.0
1	A	112(B)	SER	2.0
2	K	59	ASP	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	SO4	G	301	5/5	0.75	0.19	116,117,121,124	0
4	SO4	D	301	5/5	0.79	0.17	78,104,108,110	0
4	SO4	K	302	5/5	0.81	0.16	109,111,111,112	0
4	SO4	K	304	5/5	0.81	0.34	127,128,129,134	0
4	SO4	K	301	5/5	0.85	0.23	103,103,107,108	0
4	SO4	B	301	5/5	0.87	0.14	118,118,120,121	0
4	SO4	E	302	5/5	0.89	0.19	114,116,117,117	0
4	SO4	E	301	5/5	0.92	0.17	90,96,98,99	0
4	SO4	K	303	5/5	0.93	0.16	98,100,102,106	0
4	SO4	J	302	5/5	0.94	0.14	72,77,88,89	0
4	SO4	J	301	5/5	0.98	0.09	40,50,60,67	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.